



TITLE:

# 回帰関数の推定における停止規則 について (マルコフ・ゲーム理論と その周辺)

AUTHOR(S):

磯貝, 英一

---

CITATION:

磯貝, 英一. 回帰関数の推定における停止規則について (マルコフ・ゲーム理論とその周辺). 数理解析研究所講究録 1982, 460: 214-228

ISSUE DATE:

1982-06

URL:

<http://hdl.handle.net/2433/103107>

RIGHT:

## 回帰関数の推定における停止規則について

新潟大 理学部 磯貝英一

### §1. 序

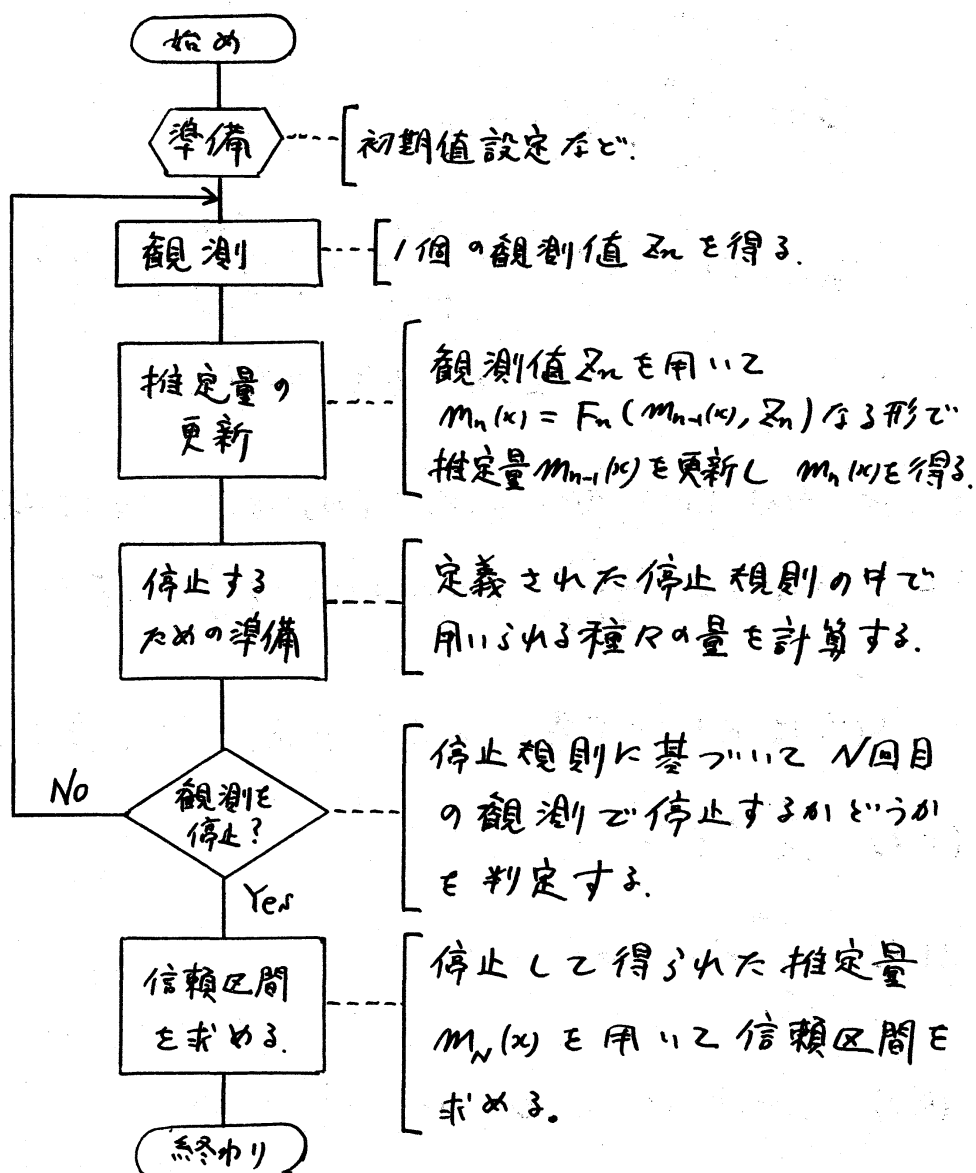
$Z = (X, Y)$ ,  $Z_1 = (X_1, Y_1)$ ,  $Z_2 = (X_2, Y_2)$ , ... は独立で同一の (ルベグ測度に関する) 確率密度関数 (p.d.f.)  $f^*(x, y)$  を持つ  $R \times R$  に値をとる確率ベクトルとする。ただし  $f^*(x, y)$  は未知とする。 $m(x) = E[Y|X=x]$  を  $X=x$  における  $Y$  の回帰関数とする。以下では  $m(x)$  は  $R$  上で有限であると仮定する。

さて、我々は観測値  $Z_1, Z_2, \dots$  を得て一点  $x$  における  $m(x)$  を区間推定したい。最近電子計算機の進歩に伴って計算上計算機に適した recursive estimators の提案されるようになっていく。(p.d.f. の推定における recursive estimators に関してはたとえば [2] を参照) しかし反復計算を行っても、実際にはこの反復をどこかで停止し、停止の結果得られた推定量を用いることになる。どこで停止した方がいいかという停止の基準は問題毎に異なるであろう。

この報告では  $m(x)$  の一つの recursive estimators  $\{m_n(x)\}$  を提案し停止規則を定義して一点  $x$  における

$m(x)$  の信頼区間を見つけることを問題とする。(p.d.f. の信頼区間を見つける問題に対してはたとえば [3] を参照) なお、信頼区間を見つけるまでの流れ図は以下のとおりである。

流れ図



## § 2. 準備

まず次のような記号を導入する。

$$f(x) = \int_{-\infty}^{\infty} f^*(x, y) dy \quad g(x) = \int_{-\infty}^{\infty} y f^*(x, y) dy$$

$$g(x) = \int_{-\infty}^{\infty} y^2 f^*(x, y) dy \quad t(x) = \int_{-\infty}^{\infty} |y|^3 f^*(x, y) dy$$

$$\psi(x) = \int_{-\infty}^{\infty} y^4 f^*(x, y) dy$$

$$\begin{aligned} \text{Var}[Y|X=x] &= \frac{g(x)}{f(x)} - \left( \frac{g(x)}{f(x)} \right)^2 \quad (f(x) > 0 \text{ a.e.}) \\ &= 0 \quad (\text{その他}) \end{aligned}$$

そこで  $f(x), g(x), g(x), t(x), \psi(x)$  は  $\mathbb{R}$  上で有限である  
とせ,  $m(x) = \frac{g(x)}{f(x)} \quad (f(x) > 0 \text{ a.e.})$  とする。

$K(x)$  は  $\mathbb{R}$  上で有界な p.d.f. であり, 次の条件を満たすとする。

$$(K1) \quad \int_{-\infty}^{\infty} x K(x) dx = 0$$

$$(K2) \quad \int_{-\infty}^{\infty} x^2 K(x) dx < \infty$$

$$(K3) \quad |x| K(x) \rightarrow 0 \text{ as } |x| \rightarrow \infty.$$

$\{h_n, n \geq 1\}$  は 0 に収束する正数列で以下 4 条件のいくつかに満たすとする:

$$(H1) \quad h_{n_0} \geq h_{n_0+1} \geq \dots \quad \text{for some positive integer } n_0$$

$$(H2) \quad n_0 h_{n_0} \leq (n_0+1) h_{n_0+1} \leq \dots \quad \text{for some positive integer } n_0$$

$$(H3) \quad \lim_{n \rightarrow \infty} n h_n = \infty$$

$$(H4) \quad \sum_{n=1}^{\infty} (n^2 h_n)^{-1} < \infty$$

$$(H5) \quad \lim_{n \rightarrow \infty} \frac{h_n}{h_{n+1}} = 1$$

$$(H6) \quad \lim_{n \rightarrow \infty} n^{1+\varepsilon} h_n^5 = 0 \quad \text{for some constant } \varepsilon > 0.$$

$\pm$   $m(x)$  の recursive estimators  $\{m_n(x), n \geq 0\}$

を次で与える:  $\forall x \in R$  に対し

$$m_0(x) \equiv 0, \quad f_0(x) \equiv c \quad (c \text{ は任意の正定数})$$

$$(2.1) \quad f_n(x) = f_{n-1}(x) + a_n \{K_n(x, X_n) - f_{n-1}(x)\}$$

$$m_n(x) = m_{n-1}(x) + a_n G(f_n(x)) \{Y_n - m_{n-1}(x)\} K_n(x, X_n)$$

$$(n=1, 2, \dots)$$

ただし

$$(2.2) \quad a_n = \frac{a}{n} \quad (0 < a \leq 1) \quad n=1, 2, \dots$$

$$G(y) = \frac{1}{y} \quad (y > 0 \text{ のとき})$$

$$= 0 \quad (\text{その他})$$

$$K_n(x, y) = \frac{1}{h_n} K\left(\frac{x-y}{h_n}\right).$$

次に停止規則を定義するために以下の量を定義する.

$$g_0(x) \equiv 0, \quad g_0(y) \equiv 0$$

$$g_n(y) = g_{n-1}(x) + a_n \{Q_n(x, Z_n) - g_{n-1}(y)\}$$

$$g_n(y) = g_{n-1}(x) + a_n \{G_n(x, Z_n) - g_{n-1}(y)\}$$

$$\text{ただし } Z_n = (X_n, Y_n),$$

$$\text{各 } z = (u, y) \in R \times R \text{ に対して}$$

$$Q_n(x, z) = y K_n(x, u), \quad G_n(x, z) = y^2 K_n(x, u)$$

$$(2.3) \quad v_n(x) = \frac{g_n(y)}{f_n(x)} - \left( \frac{g_n(x)}{f_n(x)} \right)^2 \quad (f_n(x) > 0 \text{ のとき})$$

$$= 0 \quad (\text{その他})$$

### 例 2.1

(K1), (K2), (K3) を満たす例.

$$(i) \quad K(x) = \begin{cases} \frac{1}{2} & (|x| \leq 1) \\ 0 & (|x| > 1) \end{cases} \quad (ii) \quad K(x) = \begin{cases} 1 - |x| & (|x| \leq 1) \\ 0 & (|x| > 1) \end{cases}$$

$$(iii) \quad K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (-\infty < x < \infty) \quad (iv) \quad K(x) = \frac{1}{2} e^{-|x|} \quad (-\infty < x < \infty)$$

例 2.2

(H1) から (H6) までも満たす例.

$$h_n = n^{-r} \quad \left(\frac{1}{5} < r < 1\right)$$

$\mathbb{R}$  上で定義された実数値関数  $\theta$  に対して  $C(\theta)$  は  $\theta$  の連続点の集合を表わすとする.

補題 1 ( $m_n(x)$  の強一致性)

$E[Y^2] < \infty$  を仮定する. (H4) が満たされるとき,

このとき  $f(x) > 0$  なる各点  $x \in C(f) \cap C(g) \cap C(h)$  に対して

$$\lim_{n \rightarrow \infty} m_n(x) = m(x) \quad \text{a.s.}$$

が成立する, さらに  $A \cap B$  は集合  $A$  と  $B$  の積集合を表わす.

次に  $\sqrt{nh_n} (m_n(x) - m(x))$  の漸近正規性が成立するのために満たすべき, (2.2) における  $a$  と  $\{h_n\}$  に関する 1 つの条件を定義しておく.

定義 1

次の 3 つの条件が満たされるとき, 条件  $A$  が成立するという.

$0 < a \leq 1$  なるある  $a$  に対して

$$(A1) \quad \lim_{n \rightarrow \infty} n^{1-2a} h_n = 0,$$

$$(A2) \quad \text{ある正の定数 } \beta \text{ が存在して } \lim_{n \rightarrow \infty} n^{1-2a} h_n \sum_{j=1}^n j^{2(a-1)} h_j^{-1} = \beta,$$

$$(A3) \quad \lim_{n \rightarrow \infty} (n h_n)^{\frac{3}{2}} n^{-3a} \sum_{j=1}^n j^{3(a-1)} h_j^{-2} = 0,$$

ただし  $\{h_n, n \geq 1\}$  は 0 に収束する正数列であり,  $\beta$  は  $a \in \{h_n\}$  のとり方に関係した定数である。

### 例 2.3

$h_n = n^{-r} \left( \frac{1}{2} < r < 1 \right) \quad n=1, 2, \dots \quad \frac{1-r}{2} < a \leq 1$   
 とする。このとき  $\beta = \frac{1}{2a+r-1}$  として条件 A が成立する。

$C \in R$  上の連続な関数の集合,  $C^2 \in R$  上で 2 回連続微分可能な関数の集合を表わすとする。

条件 (K4) を次で与える。

$$(K4) \quad \sup_{-\infty < x < \infty} |x| K(x) < \infty$$

例 2.1 で与えられた  $K(x)$  はすべて (K4) を満たすことがわかる。



### 補題 2 (漸近正規性)

$E[Y^2] < \infty$  を仮定する。条件 A が成立し, (H3), (H4), (H6) が満たされるとする。  $g \in C$ ,  $f, g \in C^2$  かつ  $f''(x)$ ,  $g''(x)$  は  $R$  上で有界であると仮定する。  $f(x) > 0$  かつ  $\text{Var}[Y|X=x] > 0$  なる点  $x$  を考える。このとき

(i)  $k(u)$  は  $R$  上で有界である

または

(ii)  $E[|Y|^3] < \infty$ ,  $k(u)$  は  $x$  のある近傍で有界でかつ (K4) が満たされる

のいずれか一方が成立すれば

$$\sqrt{nh_n} (m_n(x) - m(x)) \xrightarrow{\mathcal{L}} N(0, \sigma^2(x)) \text{ as } n \rightarrow \infty$$

が成立する, ただし

$$(2.4) \quad \sigma^2(x) = \alpha^2 \beta \text{Var}[Y|X=x] \int_{-\infty}^{\infty} k^2(u) du / f(x),$$

“ $\xrightarrow{\mathcal{L}}$ ” は法則収束を表わす,  $N(0, \sigma^2)$  は平均が 0 分散が  $\sigma^2$  の正規分布に従う確率変数を表わす。

### §3. 結果

この節では停止規則を定義して一点  $x$  における  $m(x)$  の, 信頼係数が  $\alpha$ , 区間の長さ  $2d$  の信頼区間を見つけることにする。

$0 < \alpha < 1, d > 0$  が任意に与えられたとする。

$D = D_\alpha > 0$  は  $\Phi(D) - \Phi(-D) = \alpha$  を満たす定数とする, ただし  $\Phi$  は標準正規分布関数である。

さて停止規則  $N(d, x)$  を次のように定義する:

$$\begin{aligned} N(d, x) &= \text{smallest integer } n \geq 1 \text{ such that} \\ (3.1) \quad (D^2 B)^{-1} d^2 n h_n &\geq \frac{v_n(x)}{f_n(x)} > 0 \text{ if such an } n \text{ exists} \\ &= +\infty \text{ otherwise,} \end{aligned}$$

ただし

$$\beta = a^2 \beta \int_{-\infty}^{\infty} K^2(u) du, \quad a \text{ は (2.2), } \beta \text{ は (A2) に依る,}$$

$v_n(x)$  は (2.3) で定義されたものである。

### 注意

$a, \{h_n\}$  従って  $\beta, K(x)$ , データ  $Z_n$  が与えられれば“停止規則を規定する条件式” (3.1) は実際に計算が可能になる。

$n(d, x)$  を次のように定義する:

$$\begin{aligned} n(d, x) &= \text{smallest integer } n \geq 1 \text{ such that} \\ (D^2 B)^{-1} d^2 n h_n &\geq \text{Var}[Y|X=x]/f(x) > 0 \text{ if such an } n \\ &\text{exists} \\ &= +\infty \text{ otherwise.} \end{aligned}$$

次に  $m(x)$  の信頼区間を次のように定義する:

$$I_{n,d}(x) = [m_n(x) - d, m_n(x) + d],$$

ただし  $m_n(x)$  は (2.1) で定義されたものである。

もし  $N(d, x) < \infty$  のとき  $N(d, x)$  回で観測を停止し

$I_{N(d,x),d}(x)$  を  $m(x)$  の信頼区間とする。このとき  
ある条件の下で

$$\lim_{d \rightarrow 0} P\{m(x) \in I_{N(d,x),d}(x)\} = \alpha$$

が成立する。Chow & Robbins [1] によればこのよ  
な性質を asymptotic consistency (of fixed-width  
sequential confidence intervals) といい。

以下では上の結果を導く。

### 補題3 (停止規則の性質)

$E[Y^2] < \infty$  を仮定する。(H3), (H4), (H5) が満たさ  
れるとする。  $f(x) > 0$  かつ  $\text{Var}[Y|X=x] > 0$  で

$x \in C(f) \cap C(g) \cap C(h)$  なる点  $x$  を考える。このとき

(i)  $\psi(u)$  は  $R$  上で有界である

また

(ii)  $E[Y^4] < \infty$ ,  $\psi(u)$  は  $x$  のある近傍で有界でかつ (K4) が  
満たされる

のいずれか一方が成立すれば、以下が成り立つ。

$$P\{N(d, x) < \infty\} = 1 \quad \text{for each } d > 0$$

$$N(d, x) \uparrow \infty \quad \text{as } d \downarrow 0 \quad \text{a.s.}$$

$$\lim_{d \rightarrow 0} \frac{N(d, x) h_{N(d, x)} d^2}{D^2 \sigma^2(x)} = 1 \quad \text{a.s.}$$

$$\lim_{d \rightarrow 0} \frac{N(d, x) h_{N(d, x)}}{n(d, x) h_{n(d, x)}} = 1 \quad \text{a.s.}$$

ここに  $\sigma^2(x)$  は (2.4) で与えられたものである。

### 命題 1

補題 3 の条件の下で

$$\lim_{d \rightarrow 0} m_{N(d, x)}(x) = m(x) \quad \text{a.s.}$$

が成立する。

(証明) 補題 1, 3 と Richter [4] の定理 1 を用いれば結果が得られる。

□

### 補題 4

$E[Y^2] < \infty$  を仮定する。(H1) から (H6) までが満たされるとする。(2.2) のある  $a$  に対して条件  $A$  が成立すると仮定する。 $f, g \in C^2$  で  $f''(x), g''(x)$  は  $R$  上で

有界であるとする。  $f(x) > 0 \Rightarrow \text{Var}[Y|X=x] > 0$  で  
 $x \in C(f)$  なる点  $x$  を考え、次の条件のいずれか一方が  
 成立するとする。

(i)  $\psi(u)$  は  $R$  上で有界である。

または

(ii)  $E[Y^4] < \infty$ ,  $\psi(u)$  は  $x$  のある近傍で有界で (K4) が  
 満たされる。

このとき、もし

$$\lim_{d \rightarrow 0} \frac{N(d, x)}{n(d, x)} = 1 \text{ in probability}$$

が成立すれば

$$\sqrt{N(d, x) h_{N(d, x)}} (m_{N(d, x)}(u) - m(u)) \xrightarrow{\mathcal{L}} N(0, \psi(x)) \text{ as } d \rightarrow 0$$

が成り立つ。

さて補題 4 を用いて主定理を与えよう。

### 定理 1

補題 4 の条件の下で

$$\lim_{d \rightarrow 0} P\{m(x) \in I_{N(d, x), d}(x)\} = \alpha$$

が成り立つ。

(証明) 補題3, 4より

$$\begin{aligned} & D d^{-1} (m_{N(d,x)}(x) - m(x)) \\ &= \sqrt{\frac{D^2 \sigma^2(x)}{N(d,x) h_{N(d,x)} d^2}} \sqrt{\frac{N(d,x) h_{N(d,x)}}{\sigma^2(x)}} (m_{N(d,x)}(x) - m(x)) \\ &\xrightarrow{\mathcal{L}} N(0,1) \text{ as } d \rightarrow 0. \end{aligned}$$

従って

$$\begin{aligned} & \lim_{d \rightarrow 0} P\{m(x) \in I_{N(d,x),d}(x)\} \\ &= \lim_{d \rightarrow 0} P\{|D d^{-1} (m_{N(d,x)}(x) - m(x))| \leq D\} \\ &= P\{|N(0,1)| \leq D\} = \alpha. \end{aligned}$$

□

### 系 1

$E[Y^2] < \infty$  を仮定する。  $g \in C$ ,  $f, g \in C^2$  かつ  $f''(x)$ ,  $g''(x)$  は  $R$  上で有界であると仮定する。

$$(3.2) \quad h_n = n^{-r} \left( \frac{1}{5} < r < 1 \right) \quad n = 1, 2, \dots$$

とおき, (2.2) の  $Q$  は

$$(3.3) \quad \frac{1-r}{2} < \alpha \leq 1$$

を満足するとする。点  $x$  は  $f(x) > 0$  かつ  $\text{Var}[Y|X=x] > 0$  を満足し、補題 4 の (i) または (ii) のいずれか一方が成立するとする。

このとき

$$\lim_{d \rightarrow 0} P\{m(x) \in I_{N(d, x), d}(b_0)\} = \alpha$$

が成り立つ。

## 系 2

$(X, Y)$  は 2 変量正規分布に従うとする、

すなわち

$$f^*(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right\}\right],$$

ここに

$$0 < \sigma_1, \sigma_2 < \infty, \quad |\rho| < 1$$

とする。

$\{h_n\}$ ,  $\alpha$  はそれぞれ (3.2), (3.3) で与えられたものと

する。

このとき すべての点  $x$  に対して

$$\lim_{d \rightarrow 0} P\{m(x) \in I_{N(d,x),d}(x)\} = \alpha$$

が成り立つ。

### 参考文献

- [1] Chow, Y.S., and Robbins, H. On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.*, 36 (1965), 457-462.
- [2] Isogai, E. Strong consistency and optimality of a sequential density estimator. *Bull. Math. Statist.*, 19 (1980), No. 1 ~ 2, 55-69.
- [3] Isogai, E. Stopping rules for sequential density estimation. *Bull. Math. Statist.*, 19 (1981), No. 3 ~ 4, 53-67.
- [4] Richter, W. Limit theorems for sequences of random variables with sequences of random indices. *Theory Probability appl.*, 10 (1965), 74-84.